communications engineering

Article

https://doi.org/10.1038/s44172-024-00253-w

# SRS-Net: a universal framework for solving stimulated Raman scattering in nonlinear fiber-optic systems by physics-informed deep learning

Check for updates

Yuchen Song[1], Min Zhang[1], Xiaotian Jiang[1], Fan Zhang[2], Cheng Ju[3], Shanguo Huang[1], Alan Pak Tao Lau[4] & Danshi Wang[1]

As a crucial nonlinear phenomenon, stimulated Raman scattering (SRS) plays multifaceted roles involved in forward and inverse problems. In fibre-optic systems, these roles range from detrimental interference that impairs optical performance to beneficial effects that enables various devices such as Raman amplifier. To obtain solutions of SRS, various numerical methods customized for different scenarios have been proposed. However, these methods are time-consuming, low-efficiency, and experience-orientated, particularly in combined scenarios consisting of both forward and inverse problems. Inspired by physics-informed neural networks, we propose SRS-Net, which combines the efficient automatic differentiation and powerful representation ability of neural networks with the regularization of SRS physical laws, to obtain universal solutions for SRS of forward, inverse, and combined problems. We showcase the intuitive solving procedure and high-speed performance of SRS-Net through extensive simulations covering different scenarios. Additionally, we validate its capabilities in experiments involving the high-fidelity modelling of a wavelength division multiplexing system spanning the C + L-band with approximately 10 THz. The versatility of the SRS-Net framework extends beyond SRS, indicating its potential as a promising universal solution in other engineering problems with nonlinear dynamics governed by partial differential equations.

Nonlinear physical systems governed by partial differential equations (PDE) are fundamental topics in scientific and engineering disciplines. On solving PDEs, there are many related challenges with a focus on addressing two types of problems: forward problems and inverse problems[1,2]. The forward problems pertain to predicting system behaviour according to the initial conditions and governing physical laws, which are important for understanding and designing systems. In contrast, the inverse problems relate to determining the initial conditions or identifying the physical laws from measurements, which are crucial for optimizing and discovering systems. The principal difference between forward and inverse problems lies in the fact that inverse problems require gradient backpropagation to update system parameters and variables. This difference often renders inverse problems more challenging to solve, and they generally require multiple customized methods for solutions, such as the case of field learning in fluid mechanics[3] and topology optimization in nanophotonics[4].

As a crucial nonlinear phenomenon in nonlinear optics, Raman scattering describes the interactions between light and the vibrational modes of molecules[5]. When a material is exposed to a high-intensity incident light, it undergoes stimulated Raman scattering (SRS), wherein a significant portion of the incident field's energy can be transferred to the scattered light, known as the Stokes wave[6]. The study of SRS holds considerable significance in fundamental scientific research as well as engineering applications of various fields, including biophysics[7], biomolecular imaging[8], spectroscopy[9], and nonlinear optics[10]. In the context of fibre-optic systems, SRS plays a vital

[1]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, 100876, China. [2]State Key Laboratory of Advanced Optical Communication System and Networks, Frontiers Science Center for Nano-optoelectronics, School of Electronics, Peking University, Beijing, 100871, China. [3]College of Electronic Information, College of Electronic Information, Qingdao University, Qingdao, 266071, China. [4]Photonics Research Institute, Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China.
e-mail: danshi_wang@bupt.edu.cn

role as nonlinear interference in ultra-wideband optical transmission and serves as the fundamental physical mechanism behind important optical devices such as Raman amplifiers and Raman lasers[6]. In these scenarios, the forward problems of SRS focus on power evolution prediction based on the initial conditions (i.e. input power profiles) and prior knowledge (i.e. the governing SRS PDE with known fibre parameters). In power evolution, the optical transmission can be unidirectional or bidirectional depending on the relative directions between signal and pump waves[11–14]. For the inverse problems, the measured input and output power profiles are known, while the objectives to be determined or optimized are the unknown fibre parameters or adjustable pump powers[15–17].

Multiple numerical methods have been put forth to address both type of problems in SRS. However, many of these methods tend to be customized, primarily providing specific solutions tailored to one type of problem within particular scenarios. For forward problems, classical numerical methods with split-step iterations are only suitable for unidirectional transmission scenarios[18]. In bidirectional transmission, methods such as the shooting algorithms are required to convert the problems into unidirectional ones, which increases the computational time and may lead to incorrect results[13,14]. When it comes to inverse problems, classical numerical methods typically lack direct support for differentiation backpropagation[19]. This deficiency necessitates the incorporation of additional iterative search algorithms. Moreover, inverse problems tend to be inherently ill-posed, resulting in a time-consuming iterative searching process[20].

A more important scenario is the combined scenario, which has intricately interconnected forward and inverse problems. In Raman amplifiers, the inverse pump power optimization is crucial for achieving a target gain spectrum, which requires iterative verifications of forward power evolution[21]. Similarly, in an ultra-wideband multi-channel transmission system, inverse parameter identification and refinement should be employed before and during high-fidelity forward power evolution prediction[22]. Addressing these combined scenarios requires a combination of classical methods, entailing extensive expert knowledge to formulate suitable solution approaches with appropriate parameters[1], which highlights the need for a universal solution framework with easy-to-implement formulations.

Deep learning methods have been proposed as promising solutions for both forward and inverse problems beyond SRS. In data-driven neural networks (NNs), forward problems can be easily solved by fitting data[23], and inverse problems can also be addressed by fitting data or utilizing an forward model with the aid of automatic differentiation (AD) in NNs[24]. A series of works utilizing data-driven NNs for the forward and inverse problem of SRS has been published. Among them, the data-driven NNs were used to predict the gain and ASE noise profiles under partial loading condition[25]. To enhance the generalization ability of forward gain prediction of the Raman amplifier, fiber characteristic was embedded into the input of data-driven NNs to construct the general model[17]. For the inverse problem of pump optimization, two data-driven NNs were used to learn the forward and backward relationship between pumps and gain profiles[16]. The pumps were first determined by the backward NNs and then refined by the forward NNs, which was experimentally evaluated on different fibers[26]. In another experimental demonstration[27], data-driven NNs were used for both span loss profiles prediction and backward Raman pump optimization. Three NN models of forward prediction in Raman amplified links were proposed[28]. In addition to two NNs for the power evolution prediction, one NN model was proposed for predicting parameters of a closed-form model formula to assist the performance estimation using the Gaussian noise (GN) model. However, the acquisition of large training datasets is often expensive and occasionally prohibitive. In this aspect, a significant challenge of data-driven NNs lies in ensuring that the collected dataset adequately covers the sample space of input profiles for forward prediction or target profiles for inverse pump optimization[29]. To address this challenge, the numerical integration of the SRS equation is utilized to replace the decoder part of an autoencoder structure[29]. In this setup, the encoder NNs can output optimized pump powers for the target gain profile under the guidance of the decoder. Additionally, a fully differentiable SRS model, enabled by NN-based nonlinear interpolation of the Raman gain coefficient, is proposed[30], but limited to the forward pumping case. Most importantly, the solutions derived through these data-driven methods lack interpretability. Additionally, the identification of physical parameters can be hindered by the challenge of establishing relationships solely from data and may suffer from the issue of non-unique mapping from outputs to input parameter sets[31]. To overcome the limitations of data-driven NNs, the physics-informed neural networks (PINNs)[32] emerged as powerful PDE solvers utilizing NNs. Different from data-driven NNs, PINNs incorporate physical laws into the loss regularization to guide the training of NNs. By combining the prior knowledge of physics and representation capabilities of NNs, PINNs have demonstrated their remarkable efficacy in solving PDEs in diverse fields, including the Navier-Stokes equations in fluid dynamics[3], the diffusion-reaction equation in chemical reactions[33], the Burgers' equation in acoustics[34], and the nonlinear Schrödinger equation (NLSE) in nonlinear optics[35,36].

Inspired by PINNs, we propose a universal solution for SRS named as SRS-Net. Formulating the PINN specifically for SRS-related problems, the proposed SRS-Net exhibits superior performance compared to classical numerical methods in the corresponding aspects. In particular, the use of AD in SRS-Net eliminates the need for step or mesh generation commonly used in classical numerical methods. Thus, the entire solver domain can be approached simultaneously under the regularization of physical laws, naturally overcoming the challenge of bidirectional transmission in forward problems. Furthermore, the physical parameters and input powers can be treated as variables in the SRS-Net, thus the AD allows direct identification of physical parameters and facilitates optimization of input powers in inverse problems. Most importantly, with the SRS-Net, combined forward and inverse problems can be effectively solved without designing specific solutions for different scenarios, alleviating the burden of carefully selecting appropriate setups for different algorithms.

In this paper, the capabilities of SRS-Net are first demonstrated in addressing basic forward problems of SRS in fibre-optic transmission. This includes scenarios ranging from continuous-wave to waveform analyses in the time-domain and multi-channel propagation in the frequency-domain, which verifies the accuracy of SRS-Net on different scales. The proposed SRS-Net is further applied to tackle forward problems of predicting power evolution under unidirectional and bidirectional transmission, which holds particular significance for Raman amplifiers. Moreover, the SRS-Net is demonstrated for solving inverse problem involving the identification of fibre physical parameters, including the frequency-dependent attenuation and Raman gain, and optimization of input pump powers within the universal framework. In addition to simulations, SRS-Net is experimentally demonstrated in a wavelength division multiplexing (WDM) system covering approximately 10 THz on C + L-band, which is expected to be the next-generation backbone information transmission infrastructure[37]. In this wideband transmission system, accurately predicting power evolution has presented a substantial challenge[38]. On this experimental setup, the inverse identification of wideband frequency-dependent fibre parameter and forward prediction of high-fidelity multichannel power evolution was simultaneously conducted. Compared with classical numerical methods, SRS-Net offers a substantial improvement in prediction speed, with a speedup of two orders of magnitude. It also demonstrates enhanced efficiency in optimization and identification, achieving more than double the convergence speed while achieving superior stability with easy-to-implement formulations. The SRS-Net methodology transcends its application in SRS, showcasing its potential extension to diverse engineering systems governed by PDEs.

## Results

### Signal propagation with SRS in fibre-optic systems

The signal propagation with SRS in fibre-optic systems can be well characterized by a set of coupled PDEs accounting for the attenuation, group velocity dispersion (GVD), Kerr nonlinearity, and simplified Raman

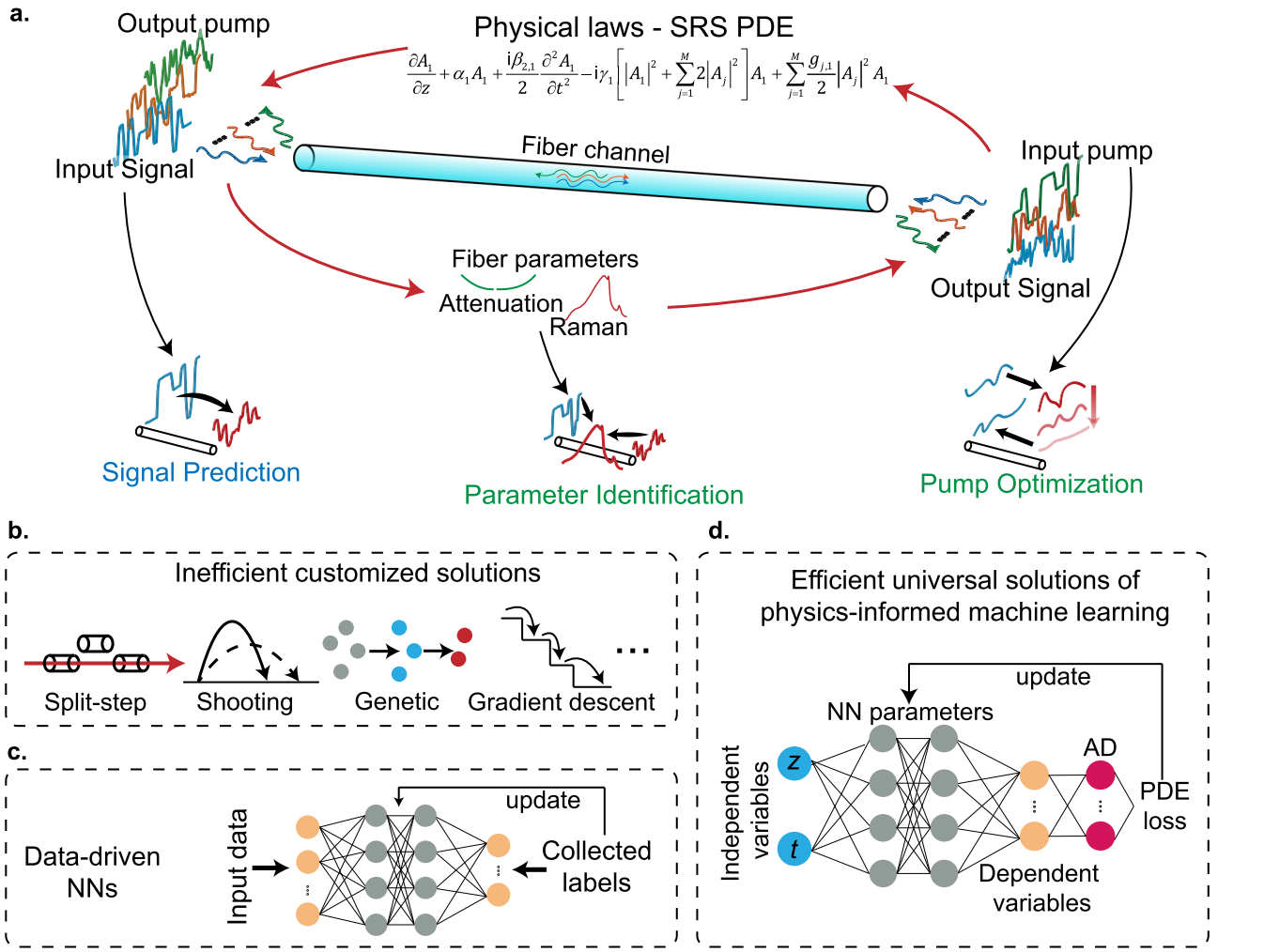



**Fig. 1 | Schematic of SRS problems in fibre-optic system and solutions. a** The complex relationships among signals, pumps, and SRS PDE with known/unknown fibre parameters that cause various SRS problems, among which we focus on one typical forward problem of signal prediction and two typical inverse problems of parameter identification and pump optimization. **b** Inefficient customized solutions using multiple classical numerical methods. **c** Data-driven NNs trained by collected labels. **d** Efficient universal solutions using physics-informed machine learning through automatic differentiation (AD).

responses[39], as expressed:

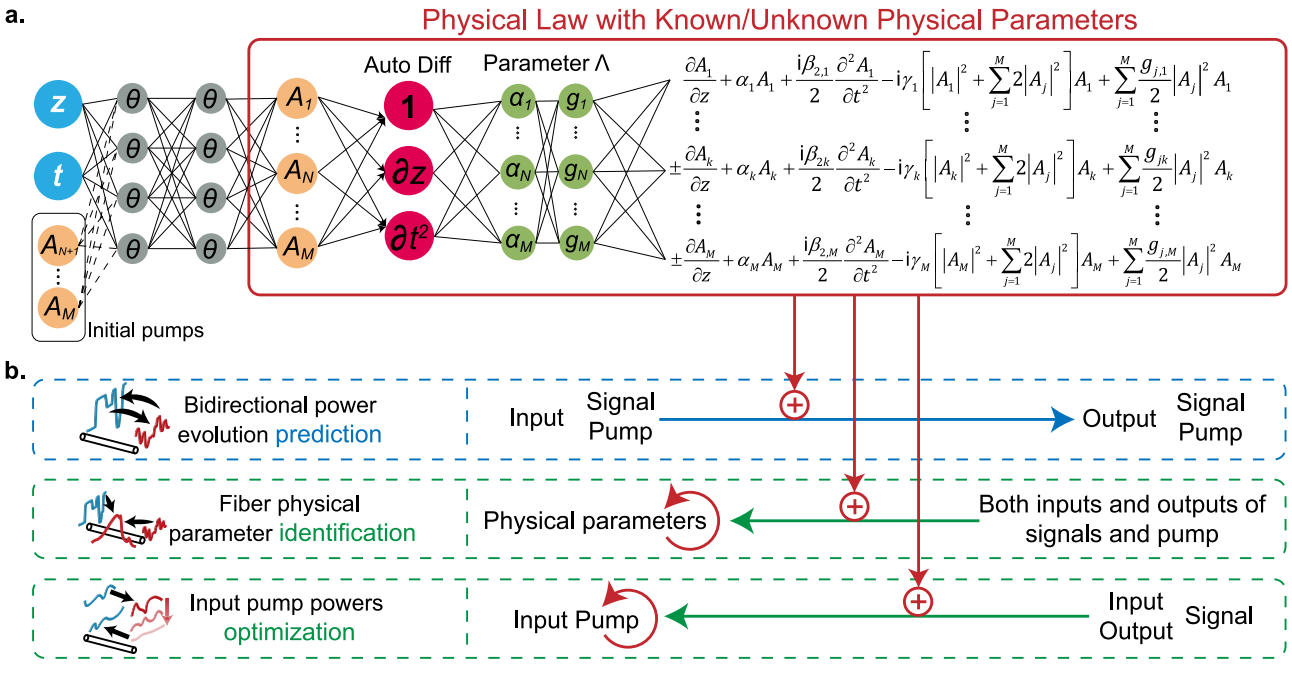
$$\pm\frac{\partial A_k}{\partial z}+\alpha_k A_k+\frac{i\beta_{2,k}}{2}\frac{\partial^2 A_k}{\partial t^2}-i\gamma_k\left[|A_k|^2+2|A_k|^2\right]A_k+\sum_{j=1}^{N}\frac{g_{jk}}{2}|A_j|^2A_k+\sum_{j=N}^{M}\frac{g_{jk}}{2}|A_j|^2A_k=0 \quad (1)$$

where $N$ represents the number of signal channels, ($M$–$N$) represents the number of pump channels, and $M$ is the total number of channels in WDM transmission. $A_k(z, t)$ is the complex-valued solution of the $k^{th}$ channel with $k = 1,2, \ldots, M$. For channels transmitted forward, the first term on the left-hand side is positive; for others transmitted backward, it is negative. Depending on the transmission direction, the initial conditions of signal and pump channels are specified at $z = 0$ or $z = z_{max}$. The symbols $\alpha_k$, $\beta_{2k}$, and $\gamma_k$ represent the coefficient of attenuation, GVD, and Kerr nonlinearity for the $k^{th}$ channel. The Raman gain coefficient $g_{jk}$ quantifies the power transfer resulting from the SRS between the $j^{th}$ and $k^{th}$ waves and is expressed as

$$g_{ik}=\pm\frac{1}{2A_{eff}}\frac{f_k}{f_j}g_R(f_j-f_k), \quad (2)$$

If $f_j > f_k$, the last term on the right-hand side of Eq. (2) is positive; otherwise, it is negative. $g_R(\Delta f)$ represents the Raman gain spectrum measured with respect to the reference frequency, and $A_{eff}$ denotes the fibre effective area.

The SRS PDE of Eq. (1) describes the full amplitude propagation in fiber transmission with SRS. However, it includes more details than necessary for scenarios with continuous wave (CW) signals and pumps, where the SRS PDE can be simplified to the coupled SRS ordinary differential equations (ODEs). The SRS PDE of Eq. (1) effectively captures the fundamental behaviours of SRS in fibre-optic systems, enabling the accurate investigation among the relationships of inputs, outputs, and physical laws with known/unknown parameters, as schematically shown in Fig. 1a. The complex relationships among them lead to various SRS problems, among which we mainly focus on one forward problem of power prediction and two inverse problems of parameter identification and pump optimization. However, to solve these problems, different formulations of customized numerical methods customized for specific solutions are required as illustrated in Fig. 1b. When tackling combined problems with both forward and inverse problems, the utilization of multiple classical methods not only introduces complexity but also leads to an inefficient iterative solution process. Moreover, these approaches often necessitate expert experience-orientated adjustments of algorithm structures and parameters[17,21,22]. In




c.

| Problem type | Problem | Required data | Target |
|---|---|---|---|
| Forward problem | Bidirectional power evolution prediction | Input signal and pump power, physical parameters | Output signal and pump power |
| Inverse problem | Fiber physical parameter identification | Input signal and pump power, output signal and pump power | Physical parameters |
| | Pump power optimization | Input and output signal power, physical parameters | Input and output pump power |

**Fig. 2 | Schematic of the universal solution of SRS-Net. a** The proposed SRS-Net composes of neural networks and physics law regularization of SRS PDE. **b** The "+" signs enclosed in a red circle indicate the embedding of physical laws in the three applications. **c** Table of three typical problems that can be solved by SRS-Net classified by required data (problem inputs) and targets (problem outputs).

addition, data-driven NNs learns the black-box model, which lacks interpretability and suffer from generalization problems.

### Architecture of SRS-Net

The PINN mainly differs from classical NN in the PDE loss function as illustrated in Fig. 1c. In the PINN, the inputs and outputs of NN corresponds to the independent and dependent variables of the governing PDEs, which are typically spatial-temporal coordinates and field amplitude, respectively. Leveraging the AD of NNs, the differential terms in the governing PDEs, which involve the differentiation of outputs with respect to inputs, can be efficiently derived. These differential terms can be utilized in constructing the PDE loss function, which serves to regularize the updating of NN parameters.

Drawing inspiration from PINNs, we propose a universal framework for solving SRS-related problems, called the SRS-Net, as depicted in Fig. 2a. First, the regularization $f$ is defined as follows:

$$f(z,t,\Lambda) = \pm\frac{\partial A_k}{\partial z} + \alpha_k A_k + \frac{i\beta_{2k}}{2}\frac{\partial^2 A_k}{\partial t^2} - i\gamma_k\left[|A_k|^2 + \sum_{j=1}^{M} 2|A_j|^2\right]A_k + \sum_{j=1}^{M}\frac{g_{jk}}{2}|A_j|^2 A_k. \quad (3)$$

The solution $A_k(z,t)$ can be learned by the NN under the regularization of $f(z, t, \Lambda)$. Different from classical PINN, we take some coefficients of SRS PDE as optimizable parameters denoted by $\Lambda$, such as $\alpha_k$ and $g_R$. Note that $A_k$ is a complex-valued signal field and is represented by $u_k$, $v_k$ in the output of the NN, where $u$ denotes the real part of $A_k$ and $v$ denotes the imaginary part. Thus, there are $2 \times M$ outputs of the NN in this setup, corresponding to a total of $M$ channels. $A_k(z,t)$ can be represented by $\boldsymbol{A} = \boldsymbol{A}(z,t) \in \mathbf{R}^{1x2M}$, with boldface representing the vector solution, and the trained NN can be denoted as $\boldsymbol{A}_{\boldsymbol{\theta}} = \boldsymbol{A}(z,t,\theta)$, where $\theta$ represents the parameters of weights and biases of the NN. Consequently, the SRS PDE regularization for the NN becomes $f(z, t, \theta, \Lambda)$, as it is also determined by the parameters of the NN. The structure of SRS-Net is illustrated in Fig. 2a.

SRS-Net can be trained by minimising the mean squared error (MSE) of both data terms $L_c$ and physical-law regularization terms $L_f$:

$$L = L_c + L_f, L_c = \frac{1}{N_c}\sum_{i=1}^{N_c}|A(D_c,\theta) - C_i|^2, L_f = \frac{1}{N_f}\sum_{i=1}^{N_f}|f(D_f,\theta,\Lambda)|^2 \quad (4)$$

The data term $L_c$ represents the constraint imposed by known or measured data at coordinates denoted by $D_c = \{z_i, t_i\}_{i=1}^{Nc}$, while the

physical-law regularization term $L_f$ enforces the regularization of $f(z, t, \Lambda, \theta)$ in the domain at random collection points denoted by $D_f = \{z_i, t_i\}_{i=1}^{N_c}$. The data term $L_c$ typically includes information of initial/boundary conditions in the forward problem or the measured outputs in the inverse problem.

The physical-law regularization $L_f$ comprises the differential terms in PDE, which can be derived by applying the chain rule to differentiate the NN's output $A_\theta$ with respect to the input coordinates (e.g. $z$ and $t$) using AD. Thus, the physical laws that govern the SRS can be integrated into loss functions. Different from the data terms, which are regularized at $z = 0$ or $z = L_{max}$, the physical-law regularization terms $L_f$ are regularized across the whole solution domain from input to output. For forward problems, the $\Lambda$ in physical laws is a priori, and the $\boldsymbol{A}$ is approached by updating the parameters of NN. In this case, the loss $L_c$ represents the initial condition and can be specified at the start or the end of fibre. Note that classical split-step numerical methods are typically discretized sequential algorithms that calculate from previous states to later ones, making them only suitable for unidirectional transmission[18]. However, SRS-Net is flexible in simultaneously obtaining solutions across the whole spatiotemporal domain. The regularization of physical laws makes SRS-Net possible to obtain solutions for both unidirectional and bidirectional transmissions, where output powers can be obtained with input powers and physical parameters, as the table shown in Fig. 2b and c. For inverse problems, the input and output powers are required for physical parameter identification. When the input and output signal powers with physical parameters are known, the input and output pump powers can be optimized. For SRS-Net, the incorporation of physical laws allows the refinement of the unknown or inaccurate physical parameters $\Lambda$ towards accurate ones with the aid of measured data. In this context, the loss $L_c$ represents the measured data, which are typically the paired inputs and outputs at $z = 0$ or $z_{max}$. In inverse problems, SRS-Net can approach the field solution $\boldsymbol{A}$ and the physical parameters $\Lambda$ simultaneously with the updated NN parameters. Ultimately, both the loss of measured data $L_c$ and physical-law regularization $L_f$ can be minimised.

## Energy transfer for continuous-wave

For the first case, we investigated a basic scenario where two waves co-propagate at two different frequencies: a signal wave at the low frequency and a pump wave at the high frequency. Under the CW assumption, the signal rate and format can be ignored. Consequently, the terms in Eq. (1) that account for dispersion and Kerr nonlinearity can be removed, resulting in the simplified version of SRS-Net as $A_\theta = A(z, \theta, \Lambda)$. With this simplification, numerical split-step techniques of simpler version should be employed to solve this set of two coupled SRS ODE with iterative steps. Obtaining solutions with SRS-Net, the input of SRS-Net is transmission distance $z$ while the output is the power on different waves. The initial conditions of both signal and pump are specified through $L_c$ at $z = 0$.

We selected launch pump powers of 20 mW, 40 mW, and 60 mW, respectively, along with launch signal powers of 2 mW, 4 mW, and 8 mW. A standard single-mode fibre (SSMF) is assumed with typical parameter values, i.e., $\alpha = 0.2$ dB/km and $g_R = 1 \times 10^{-13}$ m/W. The power evolution of the pump and signal waves over the 80 km transmission is shown in Fig. 3a,

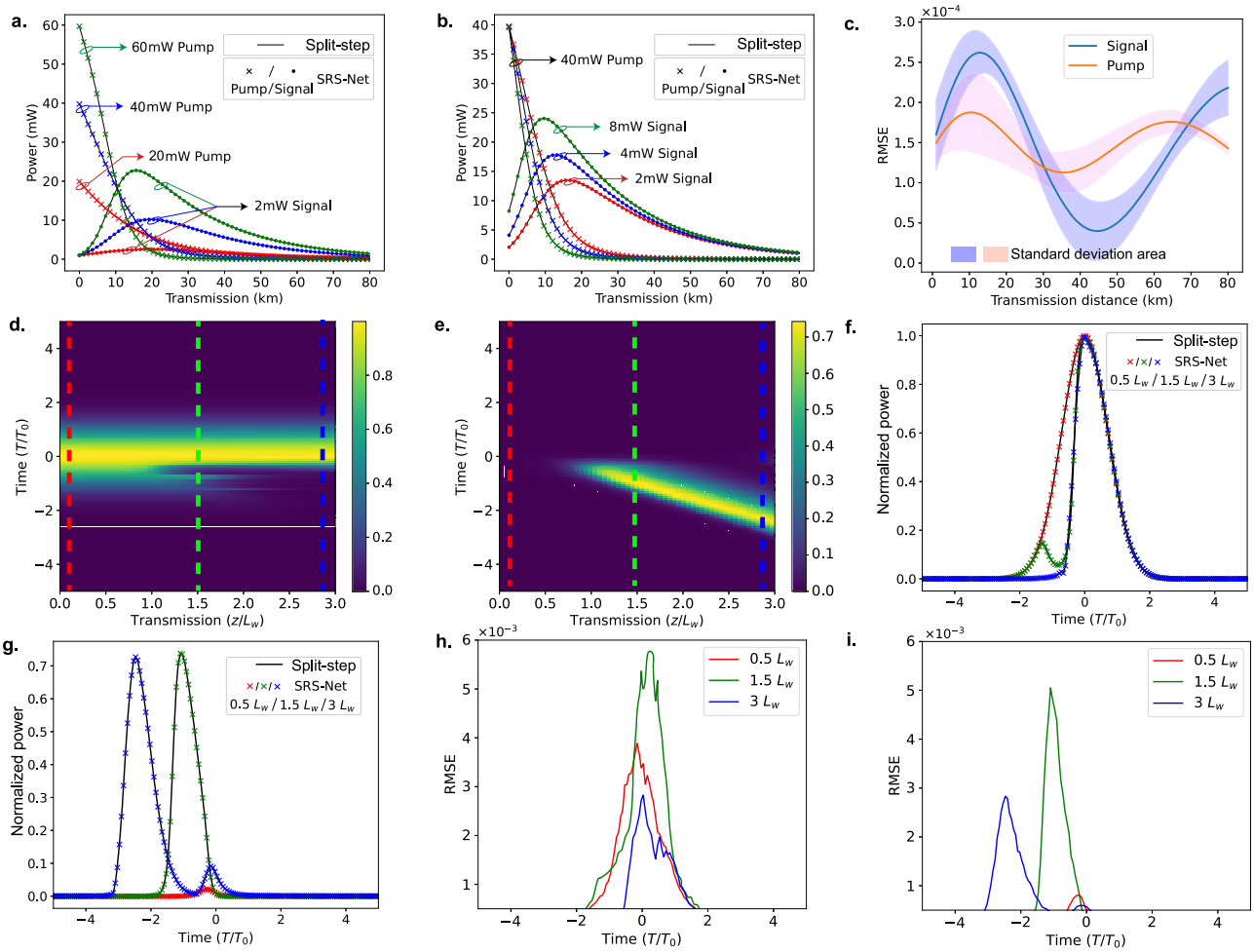



**Fig. 3 | Results of energy transfer for continuous-wave and waveform propagation on time-domain.** The energy transfer between (**a**) 2 mW launch signal and 20/40/60 mW launch pump, **b** 2/4/8 mW launch signal and 40 mW launch pump. **c** Prediction RMSE of SRS-Net for pump and signal in **a** and **b**. The time-domain waveform propagation of two pulses on the (**d**) pump wave and (**e**) signal wave, respectively, and cross-section profiles at **f** pump wave and (**g**) signal wave at z = 0.5/1.5/3 $L_W$. Prediction RMSE of SRS-Net for (h) pump and (**i**) signal at z = 0.5/1.5/3 $L_W$.

where the launch signal power is set at 2 mW while the launch pump powers are increased from 20 mW to 60 mW, and in Fig. 3b, where the launch pump power is fixed at 40 mW while the launch signal powers are increased from 2 mW to 8 mW. Considering the outputs of SRS-Net are complex-valued signal field with two output neurons $u$, $v$ representing for the real and imaginary part for each channel, the root MSE (RMSE) between the SRS-Net results $u_n$, $v_n$ and SSFM results $u_s$, $v_s$ is calculated by $\sqrt{|u_n-u_s|^2+|v_n-v_s|^2}$. In this way, both the accuracy of the amplitude and phase of the predicted signal can be evaluated by this error. The SRS-Net solutions exhibited good agreement with numerical results utilizing the small step size of 100 m, demonstrating a maximum error of less than 0.1 dB. It can be observed that the SRS-induced energy transfer occurs faster when either the pump or signal power is high. The peak power of the signal can be observed at an earlier distance with a high-power level, and then decreases due to attenuation. Furthermore, Fig. 3c illustrates the RMSE between the results obtained from SRS-Net and those from SSFM, along with the standard deviation of 12 trials. Owing to the relatively simple scenario involving only two waves co-propagating at two different frequencies, an overall high accuracy is achieved, with RMSE below $3\times10^{-4}$. It can be observed that the RMSE for both signal and pump is slightly higher at the beginning of transmission. This can be attributed to the relatively high pump power at the start, leading to strong SRS effects and resulting in relatively higher calculation complexity.

### Waveform propagation in time-domain

Without the assumption of CW, numerical methods, such as the split-step Fourier methods (SSFM)[6], are required to solve the full form of Eq. (1). For the proposed SRS-Net, the input covers both distance $z$ and time $t$, and the output is the corresponding waveform $A(z, t, \Lambda, \theta)$. We start with the example of pulse propagation. The initial pulse for both signal and pump is 50 ps width and is learned by SRS-Net at $z=0$ through $L_c$. The full regularization of $f(z, t, \Lambda, \theta)$ is applied on the entire temporal-spatial domain.

The evolutions of two pulses on the pump and signal waves for three walk-off lengths $L_w$ are shown in Fig. 3c and d; the cross-sectional profiles of the pump and signal pulses at three walk-off lengths are shown in Fig. 3e and f. The signal pulse is initially seeded at $z=0$ and is weaker than the pump pulse with 2 mW launch power by a factor of $1\times10^{-5}$. The signal pulse begins to build up after one walk-off length, and the energy transfer is almost complete by three walk-off lengths; subsequently, the pulses are separated owing to the group-velocity walk-off. In the normal GVD regime, the energy for SRS originates from the leading edge of the pump pulse, which is evident from the two-peak structure of the pump pulse at $z=1.5L_W$. As a portion of the energy from the pump pulse is transferred to the signal pulse, the pump pulse undergoes narrowing and increased asymmetry compared to its input state. The RMSE between the pulse evolution results obtained using SSFM and SRS-Net are shown in Fig. 4h and i. Overall RMSE below $6\times10^{-3}$ indicates the high accuracy of SRS-Net. The calculation error is most noticeable at transmission around $1.5L_W$, coinciding with the transmission distance of buildup of the signal wave and stronger SRS effects. The agreement between the pulse evolution results obtained using numerical methods and SRS-Net shown in Fig. 4a and b indicates the accuracy of SRS-Net.

We further conducted tests of long bit sequence propagation on the pump and signal waves rather than a single pulse, as displayed in Fig. 4a. The symbol rate was 20 GBaud with 16-Quadrature amplitude modulation (QAM) format, and the sequence length was 16 symbols resulting in 800 ps time scale. The baseline numerical results after 80 km transmission are

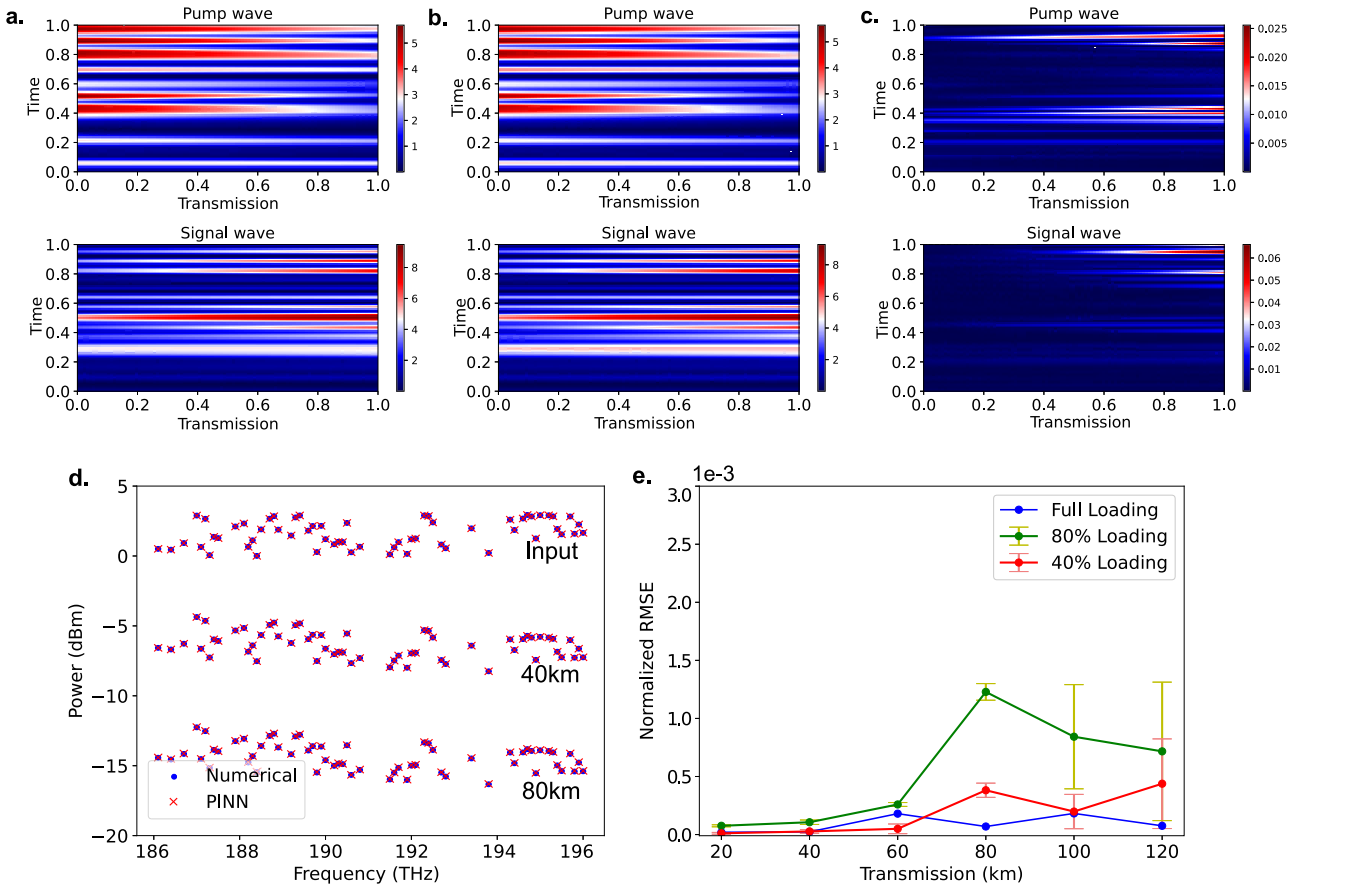



**Fig. 4 | Results of symbol sequence transmission in time-domain and multi-channel power evolution in frequency-domain. a** SRS-Net, (**b**) SSFM, and (**c**) RMSE of SRS-Net. For (**a**–**c**), x-, y-, and z-axis represent the normalized transmission distance, time scale, and optical power. **d** 60% random partial loading on 40/80 km. **e** Error distribution of full and 40/80% random partial loadings from 20 to 120 km. 12 cases are calculated for both partial loadings.

shown in Fig. 4b, and the RMSE between the SRS-Net and the numerical results are shown in Fig. 4c. The maximum RMSE was found to be smaller than 0.06, confirming the exceptional accuracy of the SRS-Net. The process of energy transfer from the pump wave to the signal wave during the transmission can be clearly observed.

### Multi-channel power evolution on frequency-domain

In fibre-optic communications of WDM systems, each channel receives energy from all high-frequency channels and transfers its energy to all low-frequency channels. In such a situation, fast and precise calculations are critical, but classical split-step methods are inefficient in calculating wide-band multi-channel transmissions[40]. Here, we selected $N = M = 96$ channels with a 100-GHz channel space propagating across the C + L band. The transmission bandwidth occupies the L-band from 186.1 THz to 190.8 THz, and the C-band from 191.4 THz to 196.1 THz, which is the default C + L-band simulation setup used in this paper. In this case, the number of outputs of SRS-Net is large, which requires the matrix of regularization of $f(z, t, \Lambda, \theta)$ to be solved in the SRS-Net.

For partial loading depicted in Fig. 4d, 60 out of 96 channels were randomly selected, and for each channel, the launch power was randomly set from 1 mW to 2 mW. The corresponding results for 40 km and 80 km transmission distances are also displayed. The input channel power spectrum, which is the only label data required for SRS-Net, is also presented, and the matching results suggest that the initial conditions are successfully learned. A detailed evaluation of the accuracy of SRS-Net solution is presented in Fig. 4e, depicting the error distributions for 40% and 80% loading, and full loading, spanning from 20 km to 120 km. The root MSE (RMSE) was used for error estimation of 12 random channel configurations for both 40% and 80% loadings. It can be observed that the standard deviation of the error slightly increases with a longer distance. This can be attributed to the fact that longer transmission will introduce more distortions and increase the difficulty of solving. However, all the RMSE are below $2 \times 10^{-3}$, which is a satisfactory level of accuracy.

### Bidirectional power evolution prediction in forward problems

In fibre-optic system, it is well known that the SRS can be leveraged for optical amplification, known as Raman amplification. During the amplification, the power of pump waves on high frequency is transferred to signal waves on low frequency. There are three common pump conditions classified by the propagation directions of the pump, including the case where the signal and pump co-propagate (CO) from one end to another and the case where the signal and pump counter-propagate (CT) from both ends. In the bi-direction (BI) condition, the input pumps are assigned at both end of the fibre and transmitted forward and backward, whereas the powers of initial signal were specified at the beginning of the fibre and transmitted forward. For scenario of CT and BI condition, including bidirectional transmission signals and pumps, shooting algorithm combined with split-step methods are generally required[13,14]. Whereas in the SRS-Net, the physical law regularization can be satisfied simultaneously on the entire domain, which naturally allows the solving of bidirectional transmission.

We setup a simulation for 100 km fibre transmission with Raman pumps under the above three conditions, and the frequencies of pumps are summarized in Table 1. The default C + L-band simulation setup covering approximately 10 THz bandwidth is used. For the SRS-Net, the outputs are the signal power of all $N$ channels combined with the pump power at ($M$-$N$) higher frequencies. For the loss function, the initial conditions for pumps and signals are assigned at $z = z_{max}$ and $z = 0$, respectively. We simultaneously enforced the physical law regularization across the entire spatial-frequency domain with known fibre parameter $\Lambda$.

The classical numerical solution of bidirectional transmission can be unstable and fail to converge, especially with high pump power levels[11]. The failure can be attributed to the inherent defects of shooting algorithms, where the conversion from bidirectional transmission to unidirectional ones is an approximation process. Here we try to obtain the solution using SRS-Net in the BI condition with a total power of pump is 400 mW. As shown in the plain

**Table 1 | Pump configuration of pump conditions of co-propagate (CO), counter-propagate (CT), and bi-direction (BI) condition**

| | CO | CT | BI |
|---|---|---|---|
| Forward pump frequency (THz) | 201/207 | - | 201/207 |
| Backward pump frequency (THz) | - | 201/205/207/208 | 201/205/207/208 |
| Pump powers (mW) | 400 | 250 | 200 |

training loss of Fig. 5a, these losses do not converge. However, for the SRS-Net, this failure can be attributed to the inappropriate initialization of NN parameters and can be easily solved by initializing the NN parameters to an appropriate point. To realize this, we conduct the adaptive weight training on this BI condition, and the training losses are illustrated in Fig. 5(a). The training is divided into three stages with different loss weights. First, the weight of pump loss is reduced, and SRS-Net is solely trained under the constraints of the input signal at $z = 0$ and the physical laws for $2 \times 10^3$ iterations. In this step, only the loss of the initial signal condition decreases because of its large weight. This leads the training of the NN to an appropriate starting point. Subsequently, for another $2 \times 10^3$ iterations in the second stage, the weight on the initial pump loss is gradually increased, and the pump loss can continue to converge. Finally, at the third stage, the weights of the initial conditions of signals and pumps are equal, and the SRS-Net keeps training until convergence.

Using the adaptive training strategy, the SRS-Net is demonstrated to effectively solve the three cases under pump conditions of CO, CT, and BI, and the longitudinal power profiles are shown in Fig. 5b-d, with a maximum error <0.3 dB compared to the ground truths. In[17], it is reported that the maximum prediction error of predicted gain profiles of a 100 km fiber was around 0.22 dB; while in ref. 28, the trained generative model can predict loss profile with RMSE less than 0.51 dB for 95% of the predictions. In comparison, the proposed SRS-Net achieves similar performance without using any labelled data.

### Fibre physical parameter identification in inverse problems

The task of parameter identification is relevant to a broad topic of scientific discovery and physical system monitoring[41]. In fibre-optic systems, the frequency-dependent attenuation and Raman gain spectrum are the two most important physical parameters that impact the power evolution. Thus, it is important to identify their accurate values from the measurements, which typically are pairs of input and output signal spectra measured in system. First, we demonstrate the identification of Raman gain profile. In this case, the $L_c$ loss function constrains the measured input and output signal spectra at $z = 0$ and $z = z_{max}$, and the accurate Raman gain spectrum $g_R$ is unknown. To validate the effectiveness of SRS-Net for parameter identification, we conducted a simulation with 130 channels across 13 THz based on three different types of fibres: SMF, dispersion-compensating fibre (DCF), and pure-silica-core fibre (PSCF), which possess different $g_R$ profiles. Other default frequency-independent parameters were used for these fibres. The fibre length was 80 km and the launch power for each channel was set to 0 dBm. The input and output signal spectra were collected for each type of fibre and used for $g_R$ identification.

In previous studies, genetic algorithm (GA) with a time-consuming iteration process was required for parameter identification[17]. The use of SRS-Net can significantly reduce the calculation time. Initially, and default linear $g_R$ (the $\Lambda$ here) was assigned to the PDE loss of SRS-Net. During the training, the NN parameters $\theta$ are updated to minimise the total loss. As the $g_R$ is incorrect, the PDE loss cannot be minimised properly when the corresponding initial and final conditions (the measured input and output signal spectrum) are rigorously learned. After every 100 epochs, the assumed $g_R$ was also updated to progressively converge towards its true value and thus further minimise the loss. The identified $g_R$ for these three types of fibres is presented in Fig. 6a, and the accuracy of parameter identification using SRS-Net was demonstrated with an RMSE below 0.2.

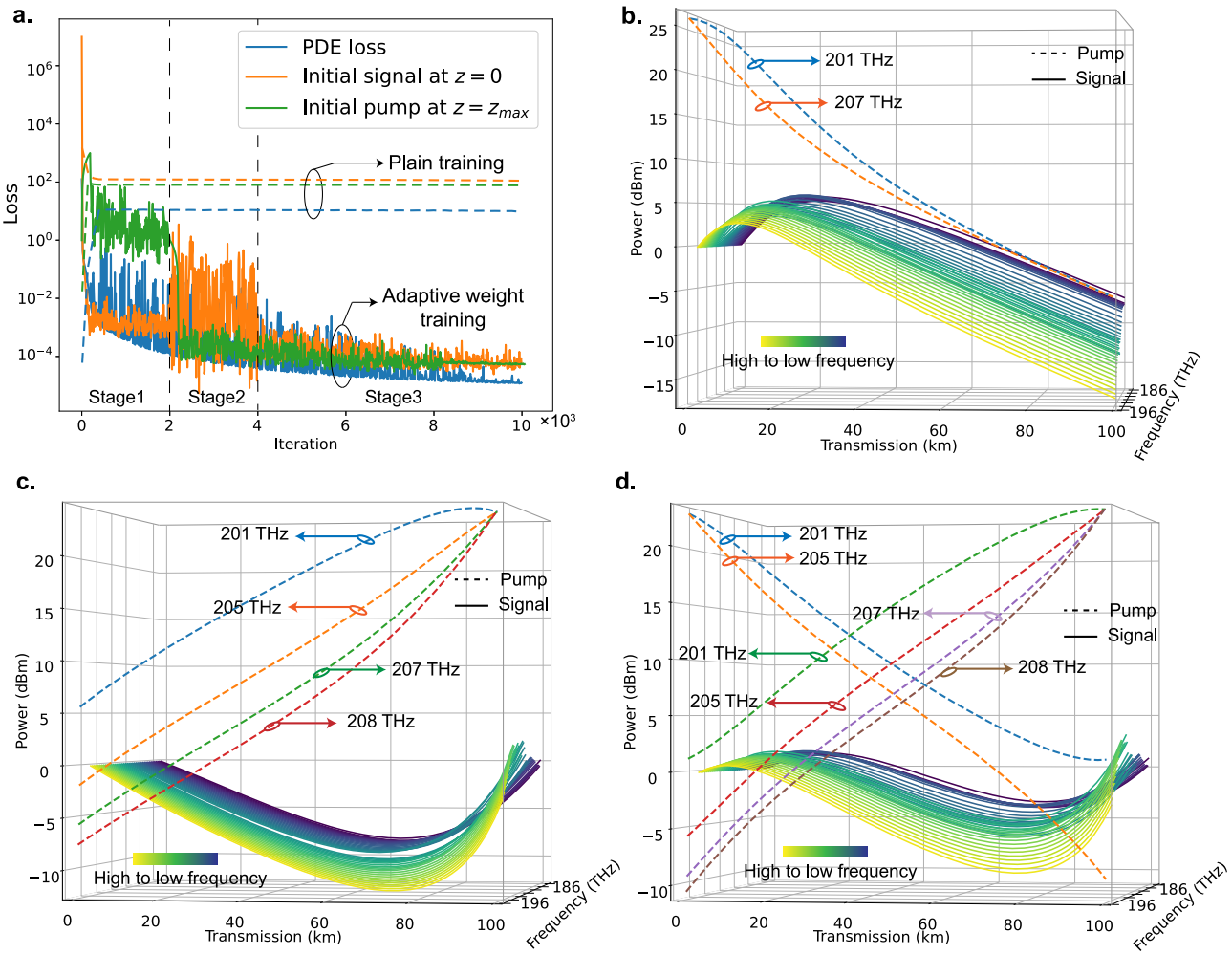



**Fig. 5 | Results of multi-channel evolution under different pump conditions.** **a** Training loss under the BI condition starting by applying on initial pump loss a small weight, which begin to increase from iteration 200. Longitudinal power profile of signal and pump wave for (**b**) co-propagate (CO), (**c**) counter-propagate (CT), and (**d**) bi-direction (BI) pump condition.

## Pump power optimization in inverse problems

In ultra-wideband fibre-optic system, flat amplified output power spectrum is often necessary for a balanced performance. The optimization of Raman pump powers to obtain a target output signal power spectrum has been a prominent subject of discussion. However, this task has proven to be an intractable challenge due to the complex nonlinear interactions between the pumps and signals. Classical prediction methods are non-differentiable, necessitating the use of iterative searching algorithms for pump power optimization[21]. Data-driven NNs have also been proposed, but massive amounts of data are typically required[15,16]. In contrast, the SRS-Net allows for efficient AD with high degrees of freedom, making it well-suited for this task.

To facilitate the optimization of pump powers, the input pump powers are used as additional NN inputs for SRS-Net as shown in the Fig. 2a. The SRS-Net outputs become $A_\theta = A(z, t, C_i, \theta)$ with $i = N, N+1, \ldots, M$. The target flat output signal spectrum is specified as $z = z_{max}$, while the input signal spectrum is specified as $z = 0$. The $\Lambda$ is known and the loss function becomes:

$$L = L_c + L_f, L_c = \frac{1}{N_c}\sum_{i=1}^{N_c}\left|A(D_c, C_j, \theta) - C_i\right|^2, L_f = \frac{1}{N_f}\sum_{i=1}^{N_f}\left|f(D_f, C_j, \theta, \Lambda)\right|^2, j = N, \ldots, M \quad (5)$$

Initially, these input pump powers are set to some default values, which are incorrect, and it becomes impossible to satisfy signal spectrum (at $z = 0$ and $z_{max}$) and the constraints of PDE simultaneously with these inappropriate input pump powers. To ensure strict adherence to these constraints, the pump power must be updated during training to minimize the total loss. Throughout the training process, where both the NN parameters $\theta$ and pump powers are updated, the pump powers can converge to optimal values while ensuring the satisfaction of PDE constraint and target signal spectrum. The regularization of the total pump power was incorporated into the loss function to avoid excessively high pump powers.

Here, we used the CT pump condition with four pumps to amplify the multi-channel signal in the same C + L band with 96 channels through 80 km transmission. The update process of the four pump powers using SRS-Net is shown in Fig. 6b, and the optimized output signal channel power spectrum is shown in Fig. 6c with the initial and target spectra. The results of the GA, utilizing the same configuration in ref. 21, are also presented for comparison. The cost function for optimization is defined as the deviation from the target spectrum. The power level of the target output signal spectrum was set slightly higher to ensure proper gain. It can be observed that SRS-Net can significantly improve the flatness with ripples reduced from 6.5 dB to 2.2 dB on the C-band and from 2.5 dB to 0.8 dB on the L-band. The optimization results of GA show 2.3 dB and 1.8 dB ripples on the C- and L-band, respectively. The flat target spectrum at 0 dBm is shown as the black line in Fig. 6c. The average difference between optimized spectrum of SRS-Net and target spectrum is

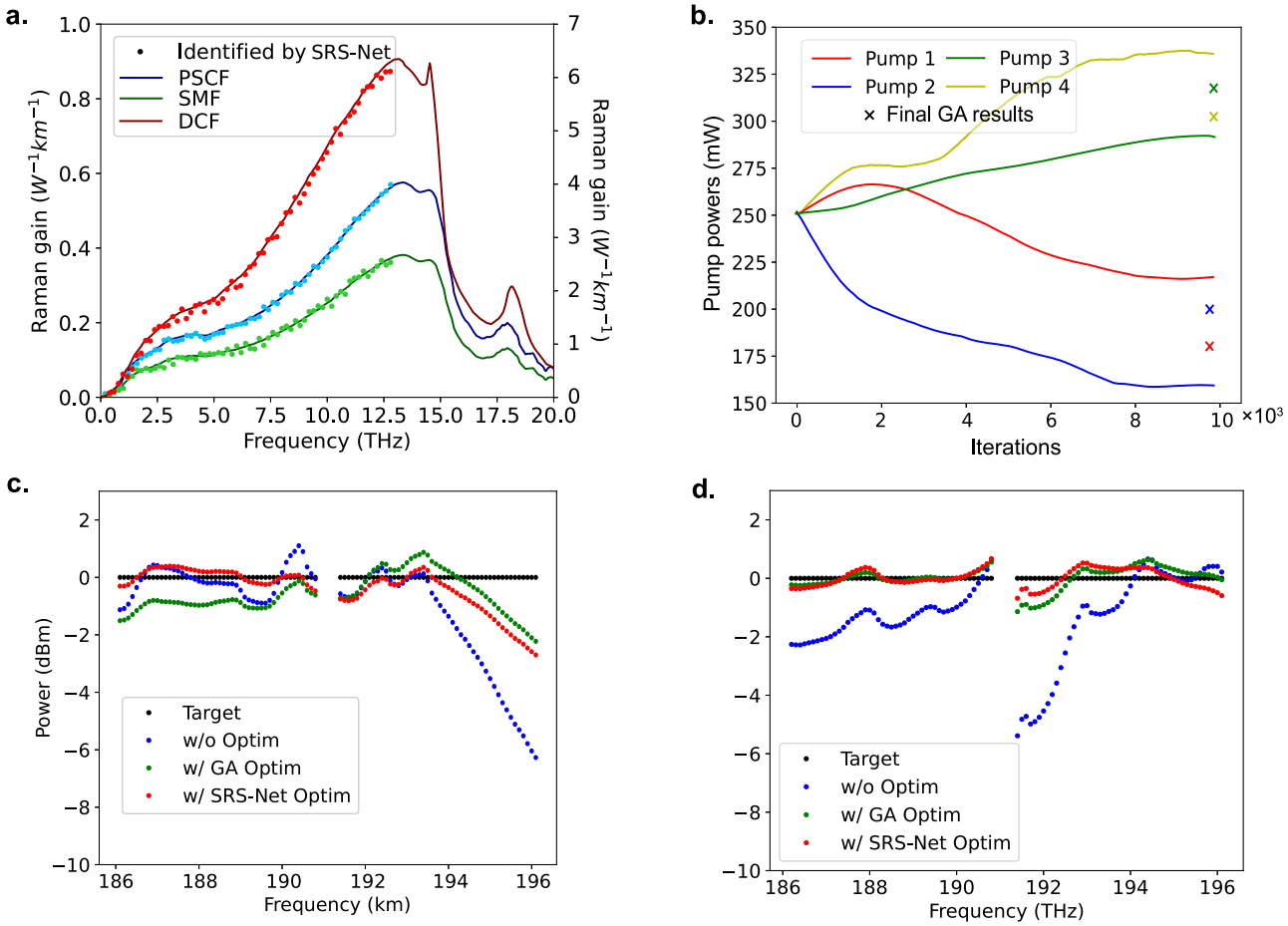



**Fig. 6 | Results of SRS-Net in inverse problems of parameter identification and pump optimization. a** Raman gain spectrum identification on three types of fibres, including single-mode fibre (SMF), pure-silica-core fibre (PSCF), and dispersion-compensating fibre (DCF). **b** Pump power optimization process by SRS-Net for four pumps under CT condition and comparison with the final GA results. Comparison of amplified power spectrum under pumps w/o optimization, w/ GA and SRS-Net optimization for (**c**) old pump frequencies and (**d**) new pump frequencies.

0.8 dB. However, this difference between optimized spectrum of GA and target spectrum is 1.7 dB, indicating the better performance of SRS-Net. However, both SRS-Net and GA fail to achieve satisfactory optimization results at the high-frequency segment of the C-band. The results presented in Fig. 6c of our original manuscript suggest insufficient pump at the high frequency of the C-band. Given that the Raman gain strength due to SRS is strongest at a frequency offset around 13 THz, the insufficient pump at high frequency of C-band can be attributed to that the selected pump frequencies are slightly low, causing most pumps to concentrate in the low-frequency part of the C-band and L-band. Consequently, achieving a flat gain profile, especially at the high frequency of the C-band, becomes challenging. Under the new pump configuration with frequencies set at 200/205/207.5/209 THz, the optimized power spectrum is shown in Fig. 6(d). It is evident that with the adoption of this new pump frequency configuration, both SRS-Net and conventional GA exhibit improved performance. The maximum deviation from the target spectrum has decreased from 5.4 dB to 0.86 dB with SRS-Net. For GA, the maximum deviation to the target spectrum is 1.3 dB, indicating satisfactory optimization performance achieved by SRS-Net.

### Universal solution of SRS-Net on experimental C + L-band fibre-optic transmission

The capability of SRS-Net has been demonstrated in the above simulations of different forward and inverse problems. Here, we set up a C + L-band WDM transmission experiment to show SRS-Net's ability in solving forward problems of power prediction and inverse problems of parameter identification in the universal framework. The WDM transmission system operating in the C + L-band is currently undergoing extensive investigation and is anticipated to serve as the next-generation information transmission infrastructure[37]. The experimental setup comprises eight spans of 75-km SSMF, except for the second span with a length of 85 km, as illustrated in Fig. 7. The transmission bandwidth occupies the L-band from 186.1 THz to 190.8 THz and the C-band from 191.4 THz to 196.1 THz with a total of 96 channels. Two commercial frequency-tuneable transponders with 400-Gb/s dual polarization (DP)-16QAM signals were deployed to measure the transmission performance for corresponding channel, and all the other 94 channels were filled with shaped amplified spontaneous emission (ASE) noise to emulate the effects of interfering channels in a cost-effective manner. The wavelength selective switch (WSS) deployed after the fifth span was not activated, and all channels experienced equal attenuation after WSS. Signals in the C- and L-bands were decoupled at the end of each span, amplified by separate C- and L-band EDFA, and then combined before launching into the next span. The connector losses for each coupler and frequency-dependent gain profiles of each EDFA were measured in advance.

Accurate power-evolution prediction is necessary for system design and performance estimation. However, for C + L-band transmission, the power distribution across the wideband is severally imbalanced mainly due to the strong SRS and frequency-dependent attenuation, which makes it difficult to obtain accurate transmitted power spectrum. Here, we demonstrate the use of SRS-Net to first identify the frequency-dependent attenuation and Raman gain spectrum, and then predict the power evolution along the experimental link based on these identified parameters.

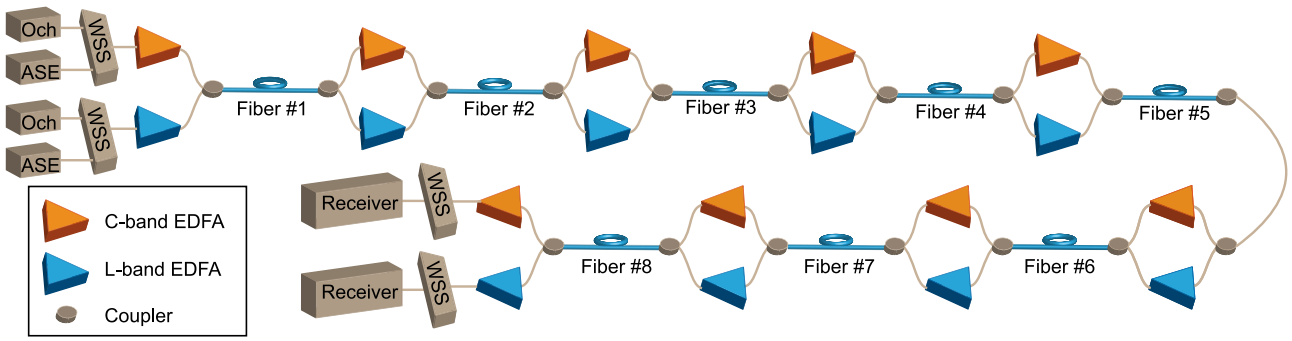



**Fig. 7 | Experimental setup for dynamic C + L-band systems.** WSS: wavelength selective switch, ASE: amplified spontaneous emission noise source, Och: 400 G DP-16QAM signal, EDFA: erbium-doped fibre amplifiers.

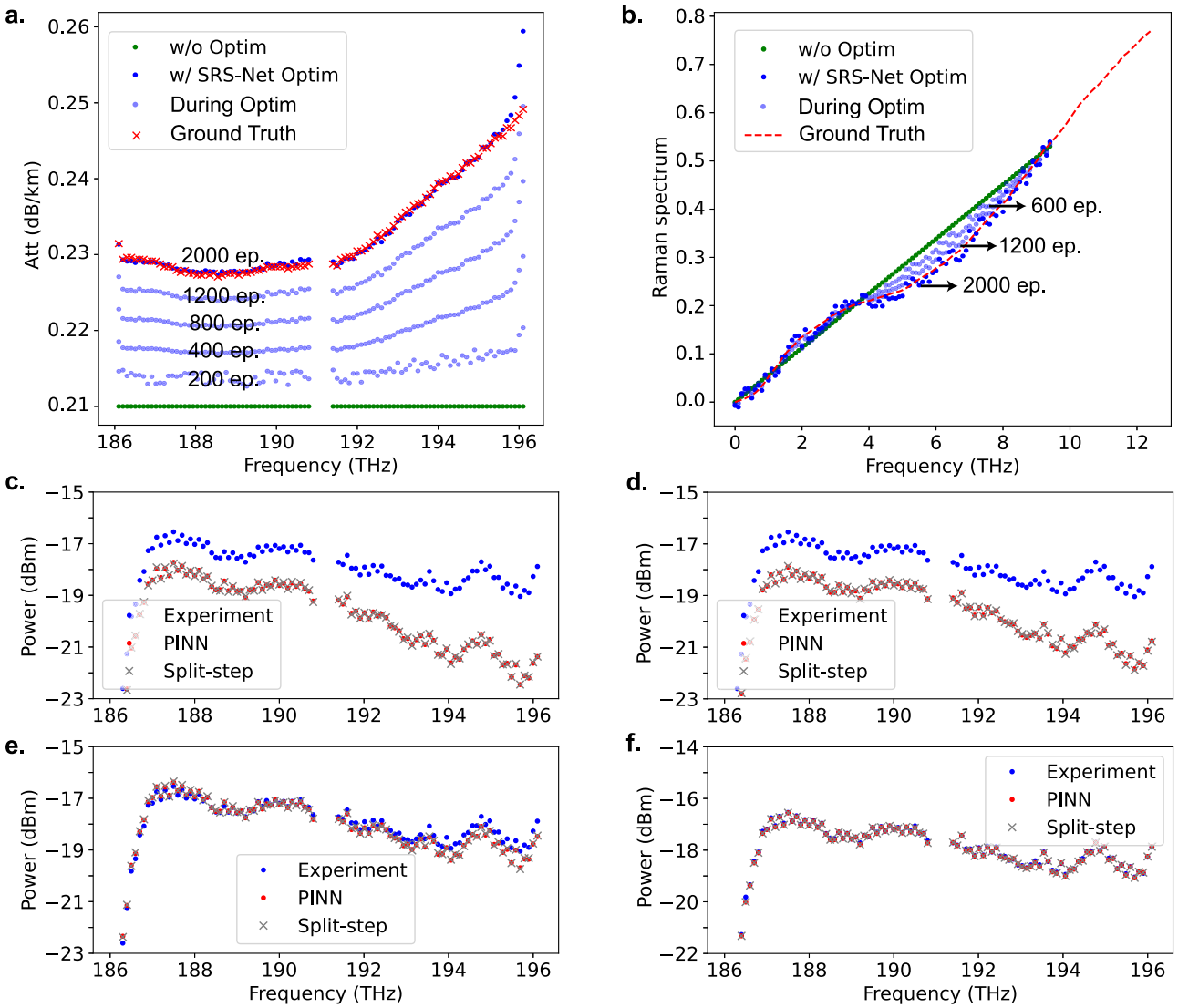



**Fig. 8 | Results of SRS-Net for parameter identification of the experimental link.** Identification of frequency-dependent (**a**) attenuation spectrum and (**b**) Raman gain spectrum for fibre#7. Visualization of the importance of accurate fibre parameters with (**c**) incorrect initial $\alpha_f$ and $g_R$, (**d**) incorrect $\alpha_f$ and correct $g_R$, (**e**) correct $\alpha_f$ and incorrect $g_R$, and (**f**) correct $\alpha_f$ and $g_R$. Split-step numerical solutions are also shown to show the accuracy of SRS-Net.

To accurately identify the $\alpha_f$ and $g_R$ of each fibre, few pairs of input and output power spectra were measured under different regular input power distributions. Here, we consider fibre#7 as an example, and only six pairs of input and output channel power spectra were used to identify $\alpha_f$ and $g_R$ simultaneously. The actual $\alpha_f$ of fiber#7 was measured using low input power in advance, and it can be observed in Fig. 8a that the initial assumed $\alpha_f$, which is constant at 0.21 dB/km, is gradually updated towards the actual $\alpha_f$, till coinciding well with the actual one in the end. To optimize $g_R$, the initially assumed linear function with respect to frequency was optimized to the actual ones with black dash line, as shown in Fig. 8b. To visualise the

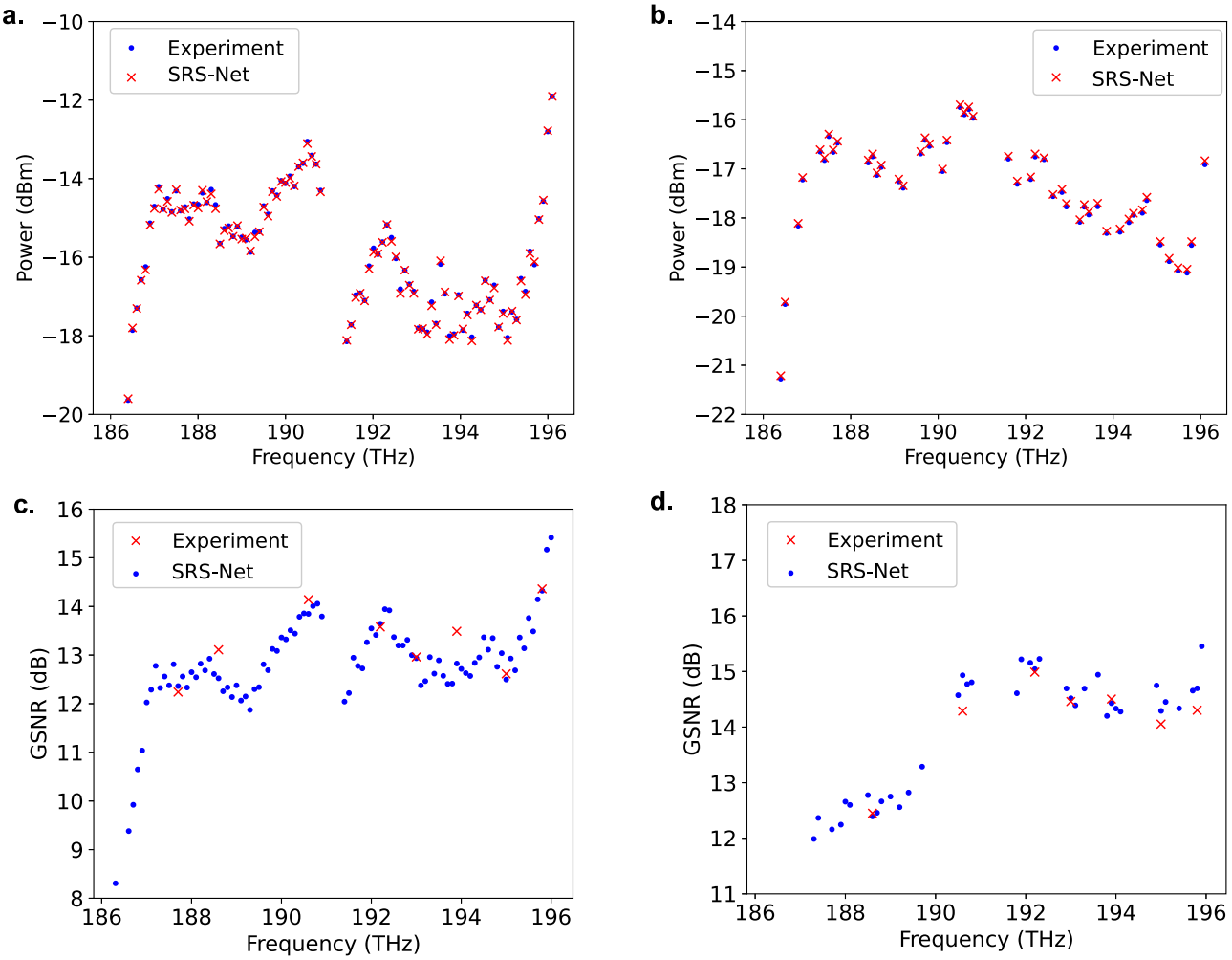



**Fig. 9 | Power and GSNR results after multi-span transmission.** Power evolution prediction of (**a**) full and (**b**) partial loading conditions. GSNR prediction of (**c**) full and (**d**) partial loading conditions.

importance of the accuracy of $\alpha_f$ and $g_R$, we conducted SRS-Net to predict four different scenarios, as shown in Fig. 8. In the first one (c), the assumed incorrect initial $\alpha_f$ and $g_R$ are used; in the second one (d), the identified $g_R$ and assumed initial $\alpha_f$ are used; in the third one (e), the identified $\alpha_f$ and assumed initial $g_R$ are used; and finally (f), both identified ones are used. It can be observed in Scenarios 1 and 2 that when an incorrect initial $\alpha_f$ is used, the error is large; in Scenario 3, the maximum absolute error is approximately 1.0 dB owing to an incorrect $g_R$. Only when both the identified $\alpha_f$ and $g_R$ are used in Scenario 4, can the maximum absolute error be reduced to below 0.1 dB, proving that accurate identification of both physical parameters is essential for obtaining a high-fidelity power evolution model.

After parameter identification using SRS-Net for each fibre in the link, we obtained SRS-Net predictions for the entire link and compared the results with the experimental optical channel monitor (OCM)-measured results under the same conditions. As shown in Fig. 9a and b, SRS- Net predictions with identified $\alpha_f$ and $g_R$ for each fibre coincide well with that of experiments for various input power distribution. Even with partial loading scenarios, the maximum power deviation is below 0.3 dB. This proved that the SRS-Net predictions with the identified coefficients were sufficiently accurate under various loading conditions. With the accurate power prediction, we further tested quality of transmission (QoT) estimation for this experimental link. The widely used Gaussian noise (GN) model[42] for calculating nonlinear interference was adopted to derive the generalised signal-to-noise ratio (GSNR), which considers both linear and nonlinear distortions. We obtained the pre-forward error correction (FEC) bit error ratio (BER) for two channels in the L-band and five channels in the C-band, which were then converted into GSNR by the pre-characterized transponder according to back-to-back BER-SNR response[43]. The prediction of GSNR is shown in Fig. 9c and d for both partial and full loading condition, and the maximum error is reduced from more than 3 dB to below 0.8 dB, which is attributed to the accurate power prediction and parameter identification of SRS-Net.

Based on this experiment, high-fidelity WDM power evolution prediction (forward problem) based on fibre parameter identification (inverse problem) are simultaneously realized through the proposed universal solution SRS-Net. The SRS-Net training for identification and prediction was performed within 10 minutes on a Tesla T4 GPU. Note that we have do our best to minimize other factors that may affect the prediction accuracy, including fluctuating laser power, changing connector losses, and load-dependent EDFA gain profiles.

## Discussion

In this study, we presented SRS-Net, a deep learning method with physical constraints, for providing universal solutions to SRS under different scenarios. SRS-Net leverages the flexibility of AD in NNs to satisfy the constraints of SRS PDE and different measurements simultaneously by updating the NN parameters, SRS PDE coefficients, and pump powers. This enables the SRS-Net to handle multiple problems related to SRS, including power evolution prediction, fibre parameter identification, and pump power optimization. Previously, these tasks typically required different numerical and searching methods, which were complex, time-consuming, and unstable. SRS-Net provides a universal solution for different SRS problems

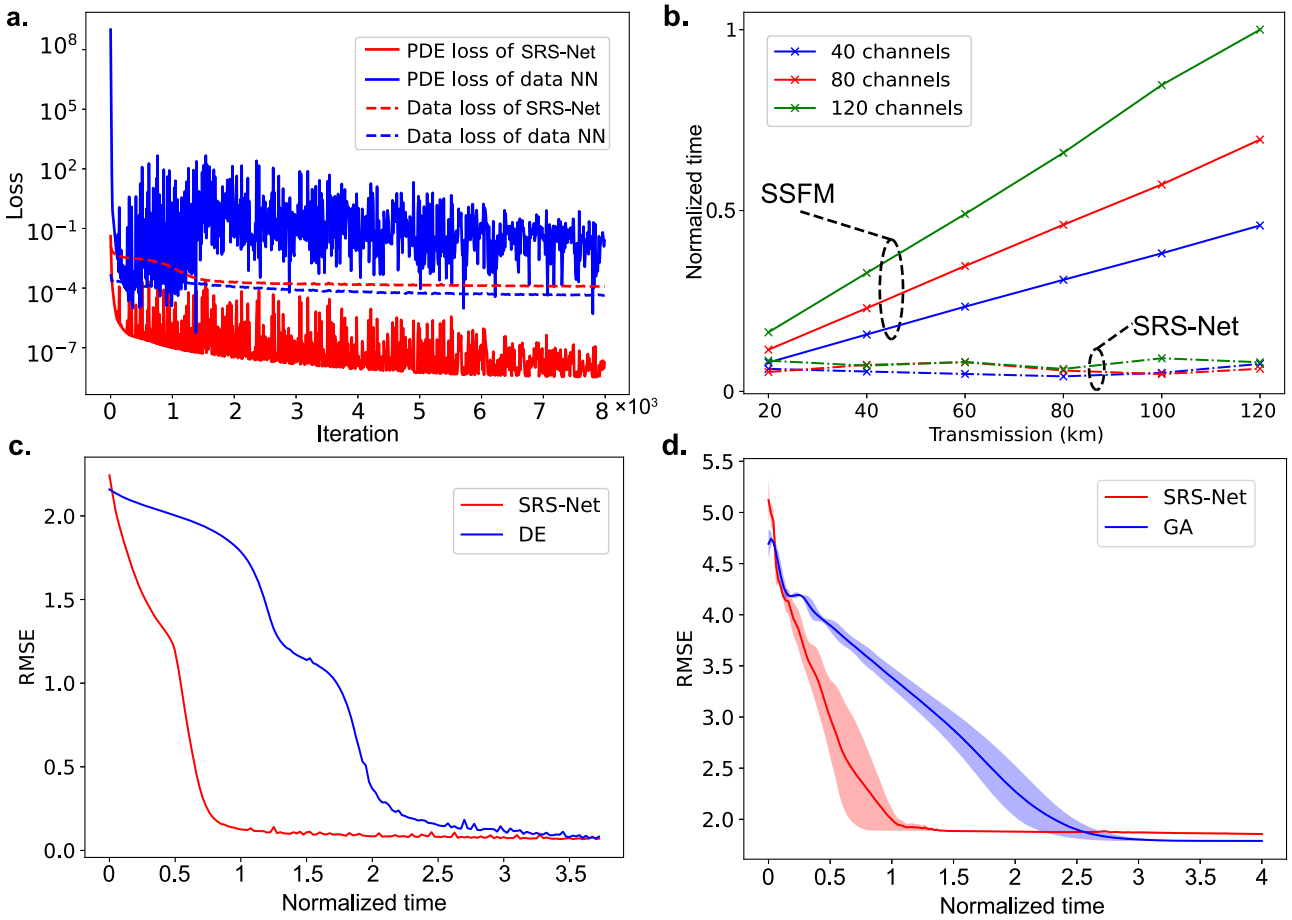



**Fig. 10 | Results of SRS-Net physics consistency, testing speed, and performance comparison. a** Loss during the training of PINN and data NN. **b** Testing time of SRS-Net and split-step Fourier method (SSFM) for 40/80/120 channels from 20 to 120 km. **c** Convergence speed of SRS-Net and DE for Raman gain spectrum identification. **d** Convergence speed of SRS-Net and GA for pump optimization.

with a consistent solution procedure and high accuracy, as demonstrated by both simulations and experiments. Moreover, SRS-Net is more flexible and faster than conventional methods. It can realise full-field multichannel power prediction in milliseconds and is insensitive to the distance and channel count owing to the parallel computation of the NNs. For parameter identification, SRS-Net differs from certain PDE coefficient discovery methods[44], which use measurement data collected from the entire spatio-temporal domain. Here, we required only the input and output data, which are significantly easier to collect in practical scenarios. The noise associated with SRS can be further incorporated into the SRS-Net to enhance the precision of its results in practical scenarios[45]. The efficacy of SRS-Net is showcased through experimental validation in a C + L-band fiber-optic transmission system, effectively addressing forward power prediction and inverse parameter identification simultaneously. The methodology of SRS-Net that combines the power of NN and physical laws has the potential to extend its applicability beyond SRS and is anticipated to address both forward and inverse problems in various other domains.

A distinctive feature of the PINN is its ability to satisfy the underlying physical laws. In this study, we demonstrate this attribute of the SRS-Net by comparing it with data-driven NNs. A classical NN is used to learn the mapping from the transmission distance to the transmitted waveforms in a data-driven manner. For the data-driven techniques, a total of 1200 pairs of 16-symbol long bit sequences are collected, out of which 1000 are used for training and the remaining 200 for testing. The symbol rate is 20 GBaud with 16-QAM format, and the maximum transmission distance is 80 km, consistent with the configuration of SRS-Net driven by physical laws. It should be noted that the SRS-Net is trained by minimizing the physical law regularization with no labelled training data while the data-driven NN (data NN) is trained by minimizing the data loss calculated from baseline labels.

The training loss is shown in Fig. 10a, and it can be observed that the SRS-physical law regularization converges to approximately $1 \times 10^{-8}$, implying that the physics law is strictly satisfied by SRS-Net. Meanwhile, the SRS-Net testing loss of predicted data between the collected labels approximates to $3 \times 10^{-4}$, which is also sufficiently accurate. For the data NN, the data loss is approximately $1 \times 10^{-4}$, showing high accuracy in the learning of transmitted waveforms; while the SRS-physical law regularization was relatively high and unstable, suggesting that the data NN failed to learn the underlying physical laws from these data. Through the incorporation of physics, the proposed SRS-Net can effectively constrain itself to a lower-dimensional manifold, enabling generalization to unseen inputs with high accuracy[1]. However, using data-driven models for extrapolation without imposing any constraints based on physical knowledge can be dangerous, particularly when the underlying physical laws are not learned. Although SRS-Net cannot mathematically guarantee convergence to the correct solution during inference, it effectively trades a minor decrease in accuracy for a significant increase in computation speed when properly trained.

SRS-Net has been demonstrated to achieve satisfactory accuracy compared to other classical numerical methods for multiple tasks. As a universal solution framework, SRS-Net also exhibited superior performance with respect to calculation speed and performance stability. For forward problems, the calculation speed of power evolution is critical for a multi-channel transmission simulation with a large channel count. The calculation speeds of SRS-Net and SSFM from 40 to 120 channels from 20 to 120 km are compared in Fig. 10(b) using the same Tesla T4 GPU. It can be observed that

the calculation time required for SRS-Net is independent of the transmission distance and only slightly increases for more channels owing to the larger NN inputs and outputs. In comparison, the calculation time of the SSFM continues to grow with the increasement of both transmission distance and channel amount. When there are 120 channels and the transmission reach 120 km, the calculation time reduced by SRS-Net is approximately a factor of two orders. Note that the calculation time refers to the time cost by inference after training.

Regarding the time advantage of SRS-Net, SRS-Net can be trained offline using physics-informed deep learning techniques for the forward problem of direct propagation. The trained SRS-Net demonstrates the ability to generalize to the unseen initial conditions, as indicated in Fig. 4d and e. In practical scenarios, a large number of direct propagation predictions are often required, such as in Raman amplifier design and optical network management. In such cases where multiple inferences are necessary, SRS-Net offers fast calculation speed at the expense of relatively long training time, similar to models trained by data-driven methods s1, s2. However, conventional methods typically require iterative split-step calculations for each inference, resulting in time-consuming processes. Figure 10b compares the speed of SRS-Net with conventional methods, highlighting the speed advantage of SRS-Net in making inferences. When a large number of forward predictions are required, the speed advantage of SRS-Net is clear even with its training time. For the inverse problem of optimization applications, the time benefit of using SRS-Net is evident from Fig. 10c, d. In such cases, where optimization of pump powers or identification of parameters is conducted during training, the total optimization time of SRS-Net is less than conventional methods.

For inverse problems, the performance stability and convergence speed are critical. Particularly for fibre parameter identification, classical methods require time-consuming iterations. In a similar task for identifying the fibre Raman gain, the differential evolution (DE) algorithm, which is a derivative-free optimization method based on the concepts of GA, was used to identify the peak of the Raman gain coefficient $g_R$[17]. It is reported that the complete optimization required approximately 30 mins to run 12 parallel instances of forward simulations[17]. We applied this algorithm to the same task with similar DE configurations[17], and then compared the convergence speed with SRS-Net for $g_R$ identification, as in the previous section of fibre physical parameter identification. In particular, both methods exhibit similar performances, with an RMSE with respect to the actual $g_R$ of approximately 0.1 dB. However, the DE method was significantly slower, taking nearly 40 mins convergence, as shown in Fig. 10c. With the same computation setup, SRS-Net only requires less than 16 mins. From the normalized time comparison in Fig. 15(c), the SRS-Net is 2.5 times faster than DE method.

For pump power optimization, we compared the GA used in ref. 33 with the proposed SRS-Net, and the convergence curve is shown in Fig. 10d. It can be found that the simultaneous optimization of the four pump powers was slightly unstable. Thus, the distribution of the standard deviation of the 12 optimizations conducted by both SRS-Net and the GA is depicted. The SRS-Net converges with only approximately 15 mins, whereas the GA requires approximately 45 mins with the same computation setup. Since the RMSE is calculated in linear scale, it is observed that the GA achieves only marginally better performance than SRS-Net towards the end. This suggests the satisfactory performance of SRS-Net, coupled with its faster computational speed.

## Methods

### Automatic differentiation in the SRS-Net

AD serves as the foundational element of the proposed SRS-Net, enabling the incorporation of PDEs into its loss functions. Typically, in the training of NNs, AD is employed through the backpropagation algorithm, which computes gradients in the NN weight space to minimize the loss function. While the comprehensive exploration of AD constitutes a substantial topic within computer science, here we focus on its application in NN training and extend its utility to the context of SRS-Net.

During the training of NNs, the backpropagation algorithm corresponds to the AD in the reverse mode, wherein derivatives with respect to NN weights are computed in the second phase of a two-phase process. Herein, we consider a simple NN with only one hidden layer of two neurons (denoted by $y_1$ and $y_2$). The inputs of this NN are denoted by $x_1$ and $x_2$, and the output is denoted by $y_3$. The weights $\omega$ of this NN are shown in the following calculation graph:

$$\begin{aligned} y_1 &= x_1 \cdot \omega_1 + x_2 \cdot \omega_3 \\ y_2 &= x_1 \cdot \omega_2 + x_2 \cdot \omega_4 \\ y_3 &= y_1 \cdot \omega_5 + y_2 \cdot \omega_6 \end{aligned} \tag{6}$$

The above calculation can also be regarded as the initial phase, where the original function code of NNs is executed forward, generating intermediate variables (such as $y_1$ and $y_2$) and capturing the dependencies in the computational graph through a bookkeeping procedure. In the second phase, derivatives are calculated by propagating adjoints in reverse, from the outputs to the inputs. We are interested in computing the contribution of the change in each weight to the change in the output $y_3$. Specifically, we seek to update the NN weights to align the output $y_3$ with the target label $t$. To achieve this, we define the loss function as $E(y_3, t)$, which is a function dependent on the NN weights. By computing the derivatives of $E$ with respect to the weights, we can determine their influence and iteratively adjust them to minimize the loss function. Here, we take the weight $\omega_5$ between neuron $y_1$ and $y_3$ as an example.

$$\frac{\partial E}{\partial \omega_5} = \frac{\partial E}{\partial y_3} \cdot \frac{\partial y_3}{\partial \omega_5} = \frac{\partial E}{\partial y_3} \cdot y_1 \tag{7}$$

The term $\partial E/\partial y_3$ can be directly calculated. For example, when $E(y_3, t) = (y_3 - t)^2$, this term equal to 2 $(y_3 - t)$. For weight $\omega_1$, the chain rule can be expressed as

$$\frac{\partial E}{\partial \omega_1} = \frac{\partial E}{\partial y_3} \cdot \frac{\partial y_3}{\partial y_1} \cdot \frac{\partial y_1}{\partial \omega_1} = \frac{\partial E}{\partial y_3} \cdot \omega_5 \cdot x_1 \tag{8}$$

The preceding calculations demonstrate the standard practice of training NNs by adjusting their weights through gradients to align the NN output with the target label. This represents the most common and widely-used application of AD. In our manuscript, the SRS-Net takes the utilization of AD a step further. Within the SRS-Net framework, incorporating differential terms into the loss function necessitates the use of AD. Typically, the differential terms of PDE to be integrated into the loss functions contain derivatives of the output ($y_3$) w.r.t the input ($x_1$ and $x_2$). This process can be effectively implemented using the chain rule, which can be expressed as follows:

$$\frac{\partial E}{\partial x_1} = \frac{\partial E}{\partial y_3} \cdot \frac{\partial y_3}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial E}{\partial y_3} \cdot \frac{\partial y_3}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} = \frac{\partial E}{\partial y_3} \cdot \omega_5 \cdot \omega_1 + \frac{\partial E}{\partial y_3} \cdot \omega_6 \cdot \omega_2 \tag{9}$$

Consequently, the differential term can be directly computed and integrated into the loss function. Subsequently, the NN parameters can be further optimized to minimize this physics-informed loss function.

For the physics regularization outlined in Eq. (3), the derivatives of the output signal field $\boldsymbol{A}$ with respect to the input distance $z$ and time $t$ can be computed using the aforementioned calculation logic. In our code implementation, we utilize the JAX library for derivative calculations, where the aforementioned chain rule calculation is inherently embedded. This allows us to obtain derivatives by the "grad" function.

**Table 2 | Hyper-parameters of SRS-Net in different cases**

| Case | Trunk width | Trunk depth | Branch width | Brunch depth | Training iterations | Training time (mins) |
|---|---|---|---|---|---|---|
| Energy transfer for continuous-wave | 50 | 4 | 50 | 4 | $8 \times 10^3$ | 12 |
| Waveform propagation on time-domain | 100 | 4 | 100 | 4 | $8 \times 10^3$ | 30 |
| Multi-channel power evolution on frequency-domain | 200 | 5 | 200 | 5 | $1 \times 10^4$ | 12 |
| Bidirectional power evolution prediction | 200 | 6 | 200 | 6 | $2 \times 10^4$ | 14 |
| Fibre physical parameter identification | 200 | 6 | 200 | 6 | $2 \times 10^4$ | 16 |
| Pump power optimization | 200 | 6 | 200 | 6 | $2 \times 10^4$ | 15 |
| Experimental C + L-band fibre-optic transmission | 200 | 6 | 200 | 6 | $1.5 \times 10^4$ | 10 |

### Hyper-parameter settings

In all cases considered in this work, the branch net and the trunk net are equipped with hyperbolic tangent activation functions (Tanh). The SRS-Net is trained using Adam optimizer with default settings. A learning rate of 0.001 and a learning rate decay of 0.9 after each 3000 iterations is used. All networks are trained using a single NVIDIA Tesla T4 GPU. The architecture of SRS-Net used in different cases is summarized in Table 2.

## Data availability

The data that support the plots within this paper are available from the corresponding author upon reasonable request.

## Code availability

The code accompanying this manuscript will be made publicly available at https://github.com/IntelligentOpticalNetworkLab/Physics-informed-SRS-Net.

Received: 23 January 2024; Accepted: 19 July 2024;
Published online: 06 August 2024

## References

1. Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Revi. Phys.* https://doi.org/10.1038/s42254-021-00314-5 (2021).
2. Tarantola, A. Inverse problem theory and methods for model parameter estimation. (SIAM, 2005).
3. Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367**, 1026–1030 (2020).
4. Lu, L. et al. Physics-informed neural networks with hard constraints for inverse design. *SIAM J. Sci. Comput.* **43**, B1105–B1132 (2021).
5. Raman, C. V. & Krishnan, K. S. A new type of secondary radiation. *Nature* **121**, 501–502 (1928).
6. Agrawal, G. P. in Nonlinear Science at the Dawn of the 21st Century 195–211 (Springer, 2000).
7. Owhadi, H. & Scovel, C. Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization: From a Game Theoretic Approach to Numerical Approximation and Algorithm Design. Vol. 35 (Cambridge University Press, 2019).
8. Kuramochi, H., Takeuchi, S. & Tahara, T. Ultrafast structural evolution of photoactive yellow protein chromophore revealed by ultraviolet resonance femtosecond stimulated Raman spectroscopy. The. *J. Phys. Chem. Lett.* **3**, 2025–2029 (2012).
9. Wang, Z., Zheng, W., Hsu, C.-Y. S. & Huang, Z. Polarization-resolved hyperspectral stimulated Raman scattering microscopy for label-free biomolecular imaging of the tooth. *Appl. Phys. Lett.* **108** (2016).
10. Dhar, L., Rogers, J. A. & Nelson, K. A. Time-resolved vibrational spectroscopy in the impulsive limit. *Chem. Rev.* **94**, 157–193 (1994).
11. Hu, J., Marks, B. S., Zhang, Q. & Menyuk, C. R. Modeling backward-pumped Raman amplifiers. *JOSA B* **22**, 2083–2090 (2005).
12. Min, B., Lee, W. J. & Park, N. Efficient formulation of Raman amplifier propagation equations with average power analysis. *IEEE Photonics Technol. Lett.* **12**, 1486–1488 (2000).
13. Liu, X. & Lee, B. Effective shooting algorithm and its application to fiber amplifiers. *Opt. Express* **11**, 1452–1461 (2003).
14. Han, Q., Jiping, N., Huayong, Z. & Zhiqiang, C. Novel shooting algorithm for highly efficient analysis of fiber Raman amplifiers. *J. Lightwave Technol.* **24**, 1946–1952 (2006).
15. Zhang, Y. et al. Raman Pump Optimization for Maximizing Capacity of C+L Optical Transmission Systems. *J. Lightwave Technol.* **40**, 7814–7825 (2022).
16. Zibar, D. et al. Inverse System Design Using Machine Learning: The Raman Amplifier Case. *J. Lightwave Technol.* **38**, 736–753 (2020).
17. de Moura, U. C., Zibar, D., Margareth Rosa Brusin, A., Carena, A. & Da Ros, F. Fiber-Agnostic Machine Learning-Based Raman Amplifier Models. *J. Lightwave Technol.* **41**, 83–95 (2023).
18. Roberts, I., Kahn, J. M., Harley, J. & Boertjes, D. W. Channel Power Optimization of WDM Systems Following Gaussian Noise Nonlinearity Model in Presence of Stimulated Raman Scattering. *J. Lightwave Technol.* **35**, 5237–5249 (2017).
19. Minkov, M. et al. Inverse design of photonic crystals through automatic differentiation. *Acs Photonics* **7**, 1729–1741 (2020).
20. O'Sullivan, F. A statistical perspective on ill-posed inverse problems. Statistical science, 502-518 (1986).
21. Perlin, V. E. & Winful, H. G. Optimal design of flat-gain wide-band fiber Raman amplifiers. *J. Lightwave Technol.* **20**, 250–254 (2002).
22. Song, Y. et al. in Optical Fiber Communication Conference. Th1F. 5 (Optica Publishing Group).
23. Wang, D. et al. Data-driven optical fibre channel modelling: A deep learning approach. *J. Lightwave Technol.* **38**, 4730–4743 (2020).
24. Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. Deep learning. Vol. 1 (MIT press Cambridge, 2016).
25. Rosa Brusin, A. M., de Moura, U. C., Curri, V., Zibar, D. & Carena, A. Introducing Load Aware Neural Networks for Accurate Predictions of Raman Amplifiers. *J. Lightwave Technol.* **38**, 6481–6491 (2020).
26. de Moura, U. C., Ros, F. D., Brusin, A. M. R., Carena, A. & Zibar, D. Experimental Characterization of Raman Amplifier Optimization Through Inverse System Design. *J. Lightwave Technol.* **39**, 1162–1170 (2021).
27. Ye, X., Arnould, A., Ghazisaeidi, A., Le Gac, D. & Renaudier, J. Optical Fiber Communication Conference. W1K. 3 (Optica Publishing Group).
28. Rosa Brusin A. M. et al. ML-Based Spectral Power Profiles Prediction in Presence of ISRS for Ultra-Wideband Transmission. *J. Lightwave Technol*. **42**, 37–47, 1 Jan.1, https://doi.org/10.1109/JLT.2023.3301897 (2024).
29. Marcon, G., Galtarossa, A., Palmieri, L. & Santagiustina, M. Model-Aware Deep Learning Method for Raman Amplification in Few-Mode Fibers. *J. Lightwave Technol.* **39**, 1371–1380 (2021).
30. Yankov, M. P., Da Ros, F., de Moura, U. C., Carena, A. & Zibar, D. Flexible Raman Amplifier Optimization Based on Machine Learning-Aided Physical Stimulated Raman Scattering Model. *J. Lightwave Technol.* **41**, 508–514 (2023).
31. Liu, D., Tan, Y., Khoram, E. & Yu, Z. Training deep neural networks for the inverse design of nanophotonic structures. *ACS Photonics* **5**, 1365–1369 (2018).

32. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
33. Kim, Y., Choi, Y., Widemann, D. & Zohdi, T. A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder. *J. Comput. Phys.* **451**, 110841 (2022).
34. Meng, X., Li, Z., Zhang, D. & Karniadakis, G. E. PPINN: Parareal physics-informed neural network for time-dependent PDEs. *Comput Methods Appl. Mech. Eng.* **370**, 113250 (2020).
35. Jiang, X. et al. Physics-Informed Neural Network for Nonlinear Dynamics in Fiber Optics. *Laser & Photonics Reviews*, **16**, 2100483 (2022).
36. Wang, D. et al. Applications of physics-informed neural network for optical fiber communications. *IEEE Commun. Mag.* **60**, 32–37 (2022).
37. Hoshida, T. et al. Ultrawideband Systems and Networks: Beyond C + L-Band. Proceedings of the IEEE, 1-17, https://doi.org/10.1109/jproc.2022.3202103 (2022).
38. Cantono, M., Schmogrow, R., Newland, M., Vusirikala, V. & Hofmeister, T. Opportunities and Challenges of C+L Transmission Systems. *J. Lightwave Technol.* **38**, 1050–1060 (2020).
39. Stolen, R. H., Gordon, J. P., Tomlinson, W. & Haus, H. A. Raman response function of silica-core fibers. *JOSA B* **6**, 1159–1166 (1989).
40. D'Amico, A., Borraccini, G. & Curri, V. Introducing the Perturbative Solution of the Inter-Channel Stimulated Raman Scattering in Single-Mode Optical Fibers. *arXiv* preprint arXiv:2304.11756 (2023).
41. Wang, H. et al. Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).
42. Semrau, D., Sillekens, E., Killey, R. I. & Bayvel, P. A Modulation Format Correction Formula for the Gaussian Noise Model in the Presence of Inter-Channel Stimulated Raman Scattering. *J. Lightwave Technol.* **37**, 5122–5131 (2019).
43. Borraccini, G. et al. Experimental demonstration of partially disaggregated optical network control using the physical layer digital twin. *IEEE Transactions on Network and Service Management* (2023).
44. Chen, Z., Liu, Y. & Sun, H. Physics-informed learning of governing equations from scarce data. *Nat. Commun.* **12**, 6136 (2021).
45. Bromage, J. Raman Amplification for Fiber Communications Systems. *J. Lightwave Technol.* **22**, 79–93 (2004).


## Acknowledgements

This work is supported in part by Natural Science Foundation of China (NSFC) (No. 62171053), in part by Beijing Nova Program (No. 20230484331), in part by Research Grants Council (RGC) General Research Fund (GRF) PolyU 15225423, in part by project 1-CD8L of the Hong Kong Polytechnic University, and in part by BUPT Excellent Ph.D. Students Foundation (No. CX2022123).

## Author contributions

Y. Song and D. Wang conceived the idea of SRS-Net. Y. Song verified the proposed method and analyzed the data of all cases. X. Jiang and C. Ju verified the proposed method in the Raman pump optimization case. F. Zhang and A.P.T. Lau developed the idea of SRS-Net. Y. Song drafted initial paper, D. Wang, M. Zhang, F. Zhang, A.P.T. Lau, and S. Huang provided the major revision and all authors contributed to the final version. D. Wang supervised the project and provided guidance to all co-authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s44172-024-00253-w.

**Correspondence** and requests for materials should be addressed to Danshi Wang.

**Peer review information** *Communications Engineering* thanks Eric Sillekens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Anastasiia Vasylchenkova. A peer review file is available.

**Reprints and permissions information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024